Detection, Classification, & Identification of Objects in Cluttered Images

Steve Elgar
Washington State University
Electrical Engineering 2752
Pullman, Washington 99164-2752
elgar@eecs.wsu.edu
Voice: (509) 335-6602; FAX: (509) 335-3818

Vinod Chandran
School of Electrical and Electronic Systems Engineering
Queensland University of Technology
Brisbane, Queensland, 4001
AUSTRALIA

Award Number: N00014-96-1-0085

LONG-TERM GOALS

The overall, long-term objective is to develop and implement the methodology to detect, classify, and identify in near-real time mines and obstacles in cluttered acoustic images.

OBJECTIVES

- Improve preprocessing filters to reduce clutter
- Incorporate size and shape information in matched filters (to reduce false alarms)
- Design a classifier combining features from higher-order spectra with those from the strength (eg, size of object) and geometry (shape) of the matched filter output

APPROACH

Our pattern recognition procedure includes preprocessing to remove background noise, matched filtering to separate the image into subsections with mine-like targets, and further classification using higher-order spectral based features. False alarms are reduced by analyzing image features with a three-stage classification scheme.

A set of features are obtained from image subsections passed through a zero-mean matched FIR filter that corresponds to an approximate shape for the target. If a mine is present, the output of the matched filter contains a large positive peak. In contrast, in the absence of a mine (ie, noise only) the output of the matched filter contains low amplitude peaks and valleys. Currently, we are testing matched filters consisting of 9 X 9 and 12 X 12 kernels that are designed to detect a horizontal edge between high values arising from reflection at the mine and low values in its shadow. A minimum distance classifier is used to detect if the distance between the high, low, and maximum peak-to-peak values in the filter output fall within a specified (by training data) threshold for a mine.

Additional features of the matched filter output are formed from the sizes of the positive and negative peak regions, the horizontal and vertical distances between the maximum and minimum values, and the relative amount of the image with concentrated regions of high or low values (determined by an adaptive threshold) (called the Euler number). These features are classified by

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **1998**	2. REPORT TYPE	3. DATES COVERED **00-00-1998 to 00-00-1998**
4. TITLE AND SUBTITLE **Detection, Classification, & Identification of Objects in Cluttered Images**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Washington State University,Department of Electrical Engineering,Pullman,WA,99164-2752**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002252.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**4**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

comparison with threshold values for mines and for noise obtained during training. The relative weight of each feature can be adjusted, and a cumulative threshold for detection established.

A large amount of the information in an image is contained in the phases of its Fourier components. Consequently, the relationships between Fourier phases can be used to form additional features useful for identifying objects within the image. We are investigating features derived from integrals of higher-order spectra of images. Advantages include:

- Retention of both Fourier amplitude and phase information
- Invariance to translation, rotation, and amplification of the object within the image
- Immunity to Gaussian noise

The bispectrum $B(f_1, f_2)$ and trispectrum $T(f_1, f_2, f_3)$ of a one-dimensional process may be defined as

$$B(f_1, f_2) = X(f_1)X(f_2)X^*(f_1 + f_2)$$

$$T(f_1, f_2, f_3) = X(f_1)X(f_2)X(f_3)X^*(f_1 + f_2 + f_3)$$

where $X(f)$ is the complex Fourier coefficient at frequency f. Features for pattern recognition are obtained from the phase of the integral of each higher-order spectrum along a radial line in bi- (eg, (f_1, f_2)) or tri- (f_1, f_2, f_3) frequency space. The integrated phase is invariant to translation, rotation, DC shifting, and amplification. Different features are obtained for different radial lines.

To extend the feature extraction to two-dimensional processes, such as images, the 2-D Fourier transform is mapped onto a polar grid using the Radon transform (parallel beam projections). The Fourier transform magnitude along a radial line forms a sequence from which higher-order spectral invariants are computed, yielding the set of invariant features. The procedure is repeated for different angles. The higher-order spectral features can be classified using K-nearest neighbors, a learning vector quantizer, or an artificial neural network.

WORK COMPLETED

The higher-order spectral feature extraction algorithm and software have been extended to include trispectral, as well as bispectral features. Features from the matched filter output and from higher-order spectra-based invariants were generated for the entire Sonar 3 database (30 images for training, 30 images for testing). An optimal set of higher-order spectra-based features from the mines in the training set was selected using principal component analysis to form linear combinations of features with minimal correlation (eg, with a diagonal covariance matrix). The 136 features (8 from the matched filter output and 64 each from linear combinations of bispectral- and trispectral-based invariants) were sorted from highest to lowest quality Q, where Q is the separation between the average value of the feature for a mine-containing region of the image from the value for a region without a mine (normalized by the sum of the standard deviations of the features for regions with and without mines). Thus, although features with high Q are best for detecting the presence of a mine, including additional features with lower Q provides more information.

A K-nearest neighbor classifier was implemented and used to classify the 30 test images in Sonar 3 after training with the 30 training images. Tests with Q > 0.15, 0.10, 0.05, and 0.025 (eg, increasing numbers of features were included) were performed. The number of nearest neighbors ranged from 3 to 27, in steps of 2, and accuracy and false alarm percentages were calculated in each case.

The output from all three classifiers (minimum distance, threshold, and K-nearest neighbor) has been combined. Software has been developed that allows each image subsection to be filtered and features to be generated. The features are investigated with each classifier sequentially to test if a detection threshold is exceeded. If no mine is detected, the next (more complex) classifier is invoked. This procedure has been applied to the entire set of images in the Sonar 3 database. The software was rewritten in a modular form, allowing different filters, features, and classifiers to be included.

RESULTS

Initial tests of fully automated (but without optimization) software with the set of filters, features, and classification schemes described above detected about 75% of the mines in the Sonar3 database, with about 10% false alarms. Operator interaction results in an improvement to about 85% accuracy.

As Q decreases from 0.15 to 0.025, the number of features included in the classification increases from 25 to 108 (of a total of 136) because more features pass the lower thresholds. Accuracies and false alarms as a function of Q are shown in Figure 1 for an 11-nearest neighbor classification. The results are not significantly different for K between about 5 and 27, except there are many more false alarms for low values of K when Q is large (eg, when there are relatively few features).

The optimal number of features for this database is about 80, indicating that the higher-order based spectral features are contributing to classification accuracy because only 8 of the features are not based on bispectra or trispectra. Trispectral-based features alone do not provide better accuracy than bispectral-based features, but combining both types of features improves classification accuracy. Including more features does not improve the results greatly (eg, Figure 1), and thus near-optimal classification may be achieved with fewer features to reduce computations.

IMPACT/APPLICATIONS

Preliminary results of this study suggest higher-order spectra-based features can be used to identify and classify patterns in noisy, cluttered images.

TRANSITIONS

No transitions took place in FY98.

RELATED PROJECTS

A proposal for complimentary investigations was submitted for an Australian Research Council Large Grant for funding in 1997 (“Detection, Classification, and Identification of Embedded Objects Using Projections and Higher-Order Spectra with Application to Classification of Viruses”). It has been revised and will be resubmitted for funding in 1999. Discussions with Dr. C. A. Butman (Woods Hole Oceanographic Institution) suggest the classification techniques we are developing may also be useful for detecting and counting larvae in images obtained from plankton pumps (eg, used during the CoOP and Duck94 field experiments).

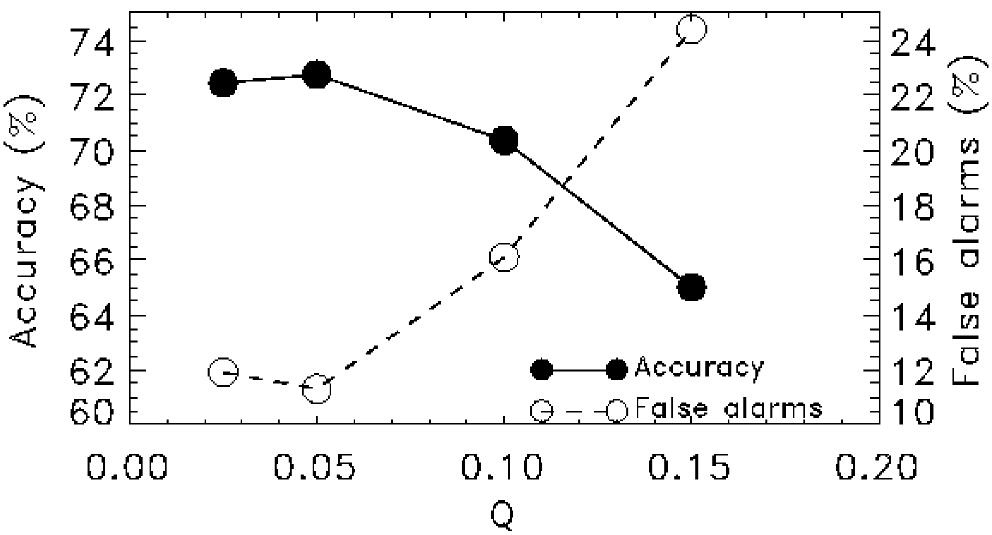

Figure 1: Percent accuracy (solid curve through closed symbols, left ordinate) and false alarms (dashed curve through open symbols, right ordinate) versus quality Q (defined in the text). The number of features used is 108, 82, 53, and 25 for Q > 0.025, 0.050, 0.10, and 0.15, respectively. Eight of the features are based on the output of a matched filter, and the rest are from higher-order spectra-based invariants. Mines were detected from the features with an 11-nearest neighbor classifier. The preliminary results shown here were obtained from fully automated software. Operator interaction increases accuracy to more than 85%.

PUBLICATIONS

V. Chandran, B. Carswell, B. Boashash, and S. Elgar, "Pattern Recognition Using Higher Order Spectral Invariants - Two Dimensional Inputs," IEEE Trans. On Image Processing **6**, 703--712, 1997.

V. Chandran, S. Slomka, M. Gollogly, and S. Elgar, "Digit Recognition Using Trispectral Features," Proc. of the ICASSP'97 Conference, Munich, pp. 3065--3068, 1997.

V. Chandran, "Instantaneous Phase of the Autocorrelation and Robustness to Signal-dependent Noise," Proc. of TENCON-97 **2**, 831--834, 1997.